Measuring adaptive control in conflict tasks

Senne Braem[1,*], **Julie M. Bugg**[2], **James R. Schmidt**[3], **Matthew J. C. Crump**[4], **Daniel H. Weissman**[5], **Wim Notebaert**[6], **Tobias Egner**[7]

[1]Vrije Universiteit Brussel, Brussels, Belgium

[2]Washington University in St. Louis, St. Louis, MO, US

[3]Université Bourgogne Franche-Comté (UBFC), Dijon, France

[4]Brooklyn College of CUNY, Brooklyn, NY, US

[5]University of Michigan, Ann Arbor, MI, US

[6]Ghent University, Ghent, Belgium

[7]Duke University, Durham, NC, US

Abstract

The past two decades have witnessed an explosion of interest in the cognitive and neural mechanisms of adaptive control processes that operate in selective attention tasks. This has spawned a large empirical literature and several theories, but also recurring identification of potential confounds and corresponding adjustments in task design to create confound-minimized metrics of adaptive control. The resultant complexity of this literature can be difficult to navigate for new researchers entering this field, leading to sub-optimal study designs. To remediate this problem, we here present a consensus view among opposing theorists that specifies how researchers can measure four hallmark indices of adaptive control (the congruency sequence effect, and the list-wide, context-specific, and item-specific proportion congruency effects) while minimizing easy-to-overlook confounds.

Keywords

Cognitive control; Conflict adaptation; Executive function; Interference effects

The quest for pure measures of adaptive control

Cognitive control (see Glossary) allows people to act in ways that are consistent with their internal goals [1]. To investigate such control, psychologists often use selective attention tasks that create **conflict** by pitting instructed task goals against incompatible stimulus information and automatic action tendencies (i.e., conflict tasks). For example, in the seminal Stroop task [2], researchers study how the ability to identify the ink color of a word varies with whether the word cues a different semantic representation and response than the

*corresponding author: senne.braem@vub.be.

color (e.g., the word BLUE in red ink; incongruent trials), or the same (e.g., the word RED in red ink; congruent trials). Participants typically respond more slowly and less accurately on incongruent trials than congruent ones. Researchers commonly consider the size of this 'congruency effect' as indicative of the signal strength of the irrelevant dimension relative to the relevant dimension, as well as the level of cognitive control applied: when congruency effects are relatively small, researchers infer there is greater recruitment of cognitive control.

Importantly, conflict tasks also allow psychologists to study modulations of congruency effects that are thought to reflect adjustments of cognitive control, which we will refer to as **adaptive control** (sometimes also called 'control learning', [3]). These dynamic adjustments of control are particularly important to measure because it is the matching of processing modes (e.g., a narrow vs. a wide focus of attention) to changing environmental demands, and/or in response to performance monitoring signals (e.g., conflict), that characterizes adaptive behavior [4]. In other words, rather than conceptualizing control as a static, time-invariant process (e.g., by assessing mean congruency effects over an entire experiment), adaptive control research is concerned with how control is regulated in a dynamic, time-varying manner. This captures both the need to deal with a changing environment, as well as the notion that control is costly and should be imposed only as much as necessary [5]. Research on adaptive control has already led to many important insights and influential theories [5–9], and continues to inspire an increasing number of studies. Moreover, beyond the basic research domain, adaptive control has been the topic of many studies and theories on developmental changes [10–12] and various clinical disorders [13–19].

However, numerous metrics of adaptive control have been put forth, criticized, and revised a number of times. Therefore, it can be difficult for new (or applied) researchers in this domain to infer what represents current best practices for studying adaptive control. In fact, many studies continue to use task designs or analysis strategies that researchers in the basic research community no longer think effectively measure adaptive control [20,21]. For example, a recent review on adaptive control in schizophrenia concluded that "there are very few clearly interpretable studies on behavioural adaptation to conflict in the literature on schizophrenia" (p. 209, [13]). Critically, systematic comparisons between older, confound-prone, and newer, confound-minimized measures of adaptive control have shown differential behavioral effects [22] and patterns of brain activity [23,24]. Therefore, the aim of the present paper is to promote best practices for investigating adaptive control, based on a current consensus view shared by different researchers in this field.

The need for "inducer" and "diagnostic" items when studying adaptive control

When employing conflict tasks, it is typically the researcher's goal to isolate changes in behavior that reflect adjustments to relatively abstract attentional settings or task representations (e.g., 'pay more attention to the target' or 'be cautious in selecting the response') as opposed to concrete settings (e.g., 'look at the green square' or 'press the left response key'). For example, in the Stroop task, abstract adjustments of control could involve paying more attention to the task-relevant color dimension or trying to inhibit the

response cued by the task-irrelevant word dimension. Such adjustments are abstract in the sense that they should lead to generalizable performance benefits that are independent of specific stimulus features or actions [25]. For instance, increased attention to the task-relevant color dimension should lead to reduced interference from the task-irrelevant word dimension regardless of the exact color and word that appears.

However, a number of researchers have pointed out that classic purported indices of adaptive control in conflict tasks can often be re-explained in terms of more basic stimulus-stimulus or stimulus-response learning processes (for a review, [21]). These considerations have led to various theoretical discussions as to how such forms of lower-level learning relate to cognitive control (Box 1). However, experts in this domain generally agree that manipulations that promote learning at this concrete level are relatively easy to avoid. Therefore, if researchers want to study adaptive control independent of low-level learning, our recommendation is that they employ paradigms that are designed to minimize opportunities for exploiting stimulus-response or stimulus-stimulus associations. We will refer to these design features for the remainder of this paper as 'confounds' (but see the section entitled When low-level learning is not a "confound").

To explain how to accomplish this goal, we brought together different researchers in this field (with different theoretical backgrounds, see Box 1) to summarize an emerging consensus view on how to design conflict tasks to study adaptive control. It quickly became apparent that the best way to summarize our view is to emphasize one key experimental design principle that enables researchers to investigate adaptive control in a confound-minimized fashion. Specifically, the principle is to distinguish between **inducer items** that trigger adaptive control and **diagnostic items** that measure the effects of adaptive control on performance.

Without going into the intricacies of what the different confounds are when investigating adaptive control, which has been extensively discussed in other papers [21,26,27], we here describe how to create tasks that avoid these confounds. We focus on four common markers of adaptive control (Box 2; Figure 1): the congruency sequence effect (CSE), the list-wide proportion congruency effect (LWPCE), the context-specific proportion congruency effect (CSPCE), and the item-specific proportion congruency effect (ISPCE). Specifically, we discuss how the inclusion of inducer and diagnostic items minimizes confounds that often prevent researchers from accurately assessing adaptive control using one of these four measures. We use the Stroop task as an example throughout this paper. However, our recommendations apply to other conflict tasks as well (Box 3).

The congruency sequence effect

The CSE, sometimes referred to as the 'Gratton effect' [28] or '**conflict adaptation** effect' [6], describes the finding that the congruency effect is reduced following incongruent compared to congruent trials (for reviews, [27,29]). The CSE is thought to measure adaptive control on a trial-by-trial basis (Box 2). However, over a decade ago, researchers noted that in typical two-alternative forced choice (2-AFC) conflict tasks with small stimulus sets, the nature of stimulus or response repetitions across consecutive trials differs for the different

conditions that researchers use to calculate the CSE. Thus, they argued that **feature integration** processes, stemming from such unequal repetitions, rather than adaptive control processes, might engender the CSE [30,31]. Later, researchers also noted that in most studies using 4-AFC tasks, which are helpful for avoiding stimulus and response repetitions, distractors are paired more often with the congruent target than with each possible incongruent target, which can lead to **contingency learning** [32]. Such tasks confound different congruency conditions with different contingency conditions and, hence, different congruency sequences with different contingency sequences. Thus, it is not possible to interpret the CSE as an index of adaptive control in such tasks (although note that some do consider contingency sequence learning a form of adaptive control, [33,34], albeit not one that involves conflict processing; see also, Box 1). Reacting to such findings, researchers began to use separate inducer items to trigger adaptive control and diagnostic items to measure the effects of adaptive control [30,35–37].

One increasingly popular way to employ inducer and diagnostic items is to create two stimulus sets (and associated response sets), each with their own congruent and incongruent items, and to alternate between these sets on a trial-by-trial basis [35–38]. For instance, in the Stroop task, one might present (a) blue and green color and word stimuli in odd trials and (b) red and yellow color and word stimuli in even trials (Figure 1A). Critically, by alternating between these two stimulus sets, the design guarantees that the inducer items that trigger adaptive control in one trial are different from the diagnostic items that measure adaptive control on the next trial, thereby ensuring that the CSE must reflect higher-level cognitive adjustments independent of stimulus-response learning. Rather than employing this alternating-sets design, one can also consider using a large set of stimuli and responses, such that each trial is a new stimulus (e.g., a picture naming task with a large set of picture-word interference stimuli; [39], Experiment 1b).

Three additional aspects deserve attention when studying the CSE. First, to avoid contingency learning confounds, one should present each unique congruent stimulus and each unique incongruent stimulus in each stimulus set equally often (e.g., the word RED in yellow should occur as often as the word RED in red; [32,40]). Thus, for example, congruent trials should occur 50% of the time when alternating between a pair of 2-AFC tasks, 33% of the time when alternating between a pair of 3-AFC tasks, and so on.

Second, while the two stimulus sets should contain different stimuli and responses, the task should be the same for both (e.g., color categorisation) as task switching reduces the CSE. That is, when using different 'stimulus sets', the idea should be to create non-overlapping sets of stimuli and responses that are all part of the same task set (e.g., name ink colors, ignore words), unless, of course, one wishes to investigate the influence of task switching on the CSE ([41–43]; for a review, [44]). Using different response sets may reduce the CSE in some situations [36,45,46], especially when using arbitrary stimulus-response mappings [47]. However, in some tasks, such as the prime-probe task, it is possible to observe robust CSEs with both arbitrary [38] and non-arbitrary [37] stimulus-response mappings, as long as participants perceive all of the stimuli and responses as belonging to the same task.

Third, one should always exclude all trials that follow an unusual event. This is because the CSE is critically dependent on what occured on the previous trial. For instance, we typically exclude all trials following an error (in addition to current-trial errors) and the first trial of each block.

Critically, in our consensus view, following the procedures described in this section will better allow researchers to attribute the CSE (Figure 2) to an adaptive control process, rather than to feature integration or contingency learning mechanisms. Note that there are still various alternative accounts as to what exactly this adaptive control process entails (Box 1), but they all involve mechanisms that go beyond low-level stimulus-response learning.

The list-wide proportion congruency effect

The LWPCE describes the finding of a smaller congruency effect in blocks of more relative to less frequent incongruent trials. The LWPCE is thought to measure global adaptations of control to the likelihood of experiencing conflict (high or low) in a particular block (list) of trials ([48], for a user's guide to the proportion congruency manipulation). In many prior studies, however, the frequency of incongruent trials was confounded with the frequency with which specific stimuli appeared. For example, a smaller congruency effect in mostly incongruent (MI) blocks than in mostly congruent (MC) blocks could reflect either adaptive control, or just having seen specific incongruent stimuli (i.e., BLUE in red) more frequently in the MI block, as participants typically respond more quickly to stimuli that appear more frequently due to low-level learning.

The best approach to studying the LWPCE while minimizing confounds also involves creating two sets of stimuli (and associated responses). The inducer items manipulate the relative frequencies of congruent and incongruent trials. In contrast, the diagnostic items measure the effect of the inducer items on the LWPCE. The diagnostic items typically consist of equal percentages of congruent and incongruent trials, although the percentages could be different, provided they remain the same in the MI and MC blocks (e.g., [49]). The inducer and diagnostic items are then randomly intermixed. As a practical example, during MI blocks, the inducer items might consist of 80% incongruent trials and 20% congruent trials, whereas the diagnostic items might consist of 50% incongruent and 50% congruent trials (Figure 1B). During MC blocks, the proportion congruency would be reversed for the inducer items (e.g., 20% incongruent, 80% congruent) but remain the same for the diagnostic items (i.e., 50% congruent).

There are two additional procedures we recommend following when investigating the LWPCE. First, to maximize the list-wide bias engendered by the inducer items in each block (i.e., such that overall PC is low in the MI block and high in the MC block), inducer items should appear more frequently than diagnostic items ([50], for sample frequencies).

Second, the inducer set should ideally be comprised of at least three different stimuli and associated responses (e.g., many studies used a Stroop task with four different words and their corresponding colors resulting in three equally frequent response options on incongruent trials; [50–52]). This is because manipulating proportion congruency on small

sets of stimuli (e.g., two colors and their corresponding words in a Stroop task) promotes contingency learning, which may weaken the triggering of adaptive control based on the proportion congruency of the list ([53,54]; but see [55,56]). This second recommendation naturally increases the number of responses in the task. For example, four colors for inducer items and two colors for diagnostic items would require six response options. These are readily accommodated in vocal versions of the Stroop task. Alternatively, one can pre-train participants on a six alternative-forced-choice (6-AFC) manual (i.e., button-pressing) conflict task.

Using such designs, we can make inferences about list-wide (global) adaptive control by analyzing the inducer and diagnostic items separately. If the congruency effect is reduced for the diagnostic items in MI blocks compared to MC blocks, this suggests that changes in control took place that cannot be attributed to the frequency with which individual items appeared (see Box 1 and Box 2, for different accounts). Finally, we note that it is also possible to investigate the LWPCE using a design wherein a unique item appears on each trial (e.g., new pictures in a picture-word Stroop), which bypasses the need for two different stimulus sets [57].

The context-specific proportion congruency effect

The CSPCE refers to the observation that the congruency effect can change when proportion congruency is manipulated across two or more contexts that vary on a trial-by-trial basis. For instance, presenting a higher proportion of incongruent stimuli in one out of two possible stimulus locations can lead to smaller congruency effects at that location than at the other location. Unlike the LWPCE, the CSPCE is thought to index adaptations to different congruency proportions within a block, which are predicted by task-irrelevant contextual features that are not part of the main task (e.g., stimulus location, [58]; color, [59]; shape surrounding the stimulus, [60]; temporal presentation windows, [61]; etc.). Relatively few studies have determined whether the CSPCE appears when measured with diagnostic items, and the evidence is mixed ([62–65]; but see [66]). Consequently, there is an ongoing controversy as to whether the CSPCE reflects adaptive control or a form of contingency learning [67]. Regardless, we believe that for researchers who want to use the CSPCE as a measure of adaptive control, the same design rules apply as for the CSE and the LWPCE.

Specifically, researchers should employ distinct sets of inducer and diagnostic items. Importantly, both sets should appear in both contexts (within the same block), but the inducer items should vary in proportion congruency (and thereby trigger adaptive control) dependent on the context. For example, in Figure 1C, the relevant color blue appears most often with an incongruent word distractor in the upper (MI) screen location but with a congruent word distractor in the lower (MC) screen location. The diagnostic items (e.g., 50% congruent in each location) measure the effects of adaptive control. Critically, if the congruency effect for the diagnostic items is smaller in the MI context (e.g., in the upper screen location) than in the MC context (e.g., in the lower screen location; Figure 2), one may conclude there is evidence for adaptive control.

The item-specific proportion congruency effect

Our final index of adaptive control is the ISPCE, which refers to the finding that the size of the congruency effect for a particular item varies with how frequently it appears in incongruent relative to congruent trials. For example, the congruency effect is smaller for target items that appear more frequently with incongruent distractors than for target items that appear more frequently with congruent distractors. Much like the LWPCE and the CSPCE, the ISPCE reflects adaptations to different proportions of congruent and incongruent trials. However, it differs from these other indices of adaptive control in that researchers manipulate the proportion congruency of different *task-relevant* items (e.g., target colors, [68]; pictures, [69,70]). The original studies aimed to manipulate the proportion congruency of the distractors [71]. However, it is preferable that the relevant stimulus feature is predictive of proportion congruency, because when the irrelevant feature is predictive (e.g., the word in the Stroop task), the ISPCE that results can be driven by contingency-learning, as shown in several studies ([69,72]; but see [69], Experiment 1; [68], Experiment 3, four-item set condition). Accordingly, when examining the ISPCE in the Stroop task, researchers manipulate proportion congruency across the different ink colors that appear randomly in each block (e.g., blue is MC, red is MI). Critically, unlike for the other three indices of adaptive control, the 'diagnostic items' should share the predictive feature (e.g., color) with the biased inducer items. Otherwise, item-specific adjustments cannot be assessed.

Design-wise, examining the ISPCE begins with the creation of two stimulus sets that vary in their task-relevant feature (e.g., color in the Stroop task). Unlike for the LWPCE and CSPCE, these stimulus sets both consist of inducer items; that is, the triggers of adaptive control. In the Stroop task, for example, one set of colors (e.g., blue and green) appears more frequently with a congruent word (MC items) while the other set of colors (e.g., red and yellow) appears more frequently with an incongruent word (MI items). During the task, these items are randomly intermixed (Figure 1D). Further, the item sets overlap, meaning, in our example, that the colors red and yellow would appear not just with the words 'red' and 'yellow', but also with the words 'blue' and 'green', and vice versa for the colors blue and green. This might discourage participants from relying on contingency learning based on the distractor word to predict the target response [69], which can occur when the sets do not overlap ([71,72], but see [68], Experiment 3, four-item set condition). One typically tests for the presence of an ISPCE by determining whether the congruency effect is smaller for MI (e.g., red and yellow) items than for MC (e.g., green and blue) items. However, although the ISPCE in this design is based on differences in proportion congruency associated with the relevant dimension, the distractor words are not completely uncorrelated with responses ([69], Experiment 2; [68], Experiments 1 & 2; [70,73]).

Therefore, consistent with our general recommendation, we encourage researchers to create a third set of diagnostic items to assess the ISPCE without item-frequency differences. There currently exist two kinds of approaches using diagnostic items, but each comes with a cautionary note. First, one can create diagnostic items by choosing novel items from the same categories as the inducer items ([69], Experiment 2; [70]). For example, in a Stroop task wherein participants name a picture category (e.g., dogs, birds) while ignoring a

superimposed word (e.g., cat), one can use a novel set of exemplars as diagnostic items (e.g., new dogs and birds that are 50% congruent). Importantly, while this approach already reduces some prominent contingency learning confounds, it could still be influenced by smaller remaining distractor-response contingencies or contingency learning at a categorical level based on the distractor word ([74,75]; for a review, [21]). No study has established that these remaining contingency learning opportunities can give rise to the ISPCE in these particular designs, but researchers should be aware of this caveat.

A second approach is to use diagnostic items that involve new distractor features (e.g., words). Crucially, this approach controls for all currently known frequency and contingency learning confounds. For example, in the Stroop task, diagnostic trials consist of the MC and MI inducer colors paired equally often with incongruent non-inducer distractor words such as 'purple' or 'black' (Figure 1D). Using this design, one can examine whether responses are faster for (1) MI inducer colors paired with non-inducer incongruent words than for (2) MC inducer colors paired with non-inducer incongruent words. Such a result indicates adaptive control in the absence of frequency and contingency learning confounds. However, analogous to the CSPC, only one study has investigated item-specific control using this particular approach ([68], Experiment 2). Therefore, while we can confidently recommend this approach, it will be important to assess its robustness in future studies.

Power analysis and design planning

Our recommendation to use inducer and diagnostic items for measuring adaptive control also comes with a warning regarding statistical power. The CSE, LWPCE, ISPCE, and CSPCE are all calculated as the difference between two congruency effects, which are themselves difference scores. Taking a difference between difference scores can increase variability [76]. Hence, if a design has sufficient power to measure a congruency effect, it might not necessarily have sufficient power to measure a difference between congruency effects. Moreover, splitting a design into inducer and diagnostic items can reduce the number of trials used to measure the effect of interest (with the exception of the CSE), thus further reducing power.

Therefore, we recommend that researchers ensure their design has enough trials and participants to achieve a desired level of statistical power for assessing adaptive control. To aid with power calculations (and determining the required sample size) for adaptive control designs, Crump and Brosowsky [77] created conflictPower (https://crumplab.github.io/conflictPower/), a free R package for estimating power using monte-carlo simulations (using the method from [64]). This approach provides power estimates for congruency effects and differences between congruency effects that depend on number of participants, effect-size, number of trials per cell, and estimates about the shape of underlying reaction time distributions using ex-Gaussian parameters.

We are sensitive to the fact that acquiring more trials or bigger sample sizes can sometimes be challenging, especially in clinical studies. Therefore, we are not recommending that patient studies be withheld for the sole reason that they might lack the statistical power for drawing firm conclusions. One way to address this general issue (that is not specific to this

literature; e.g., [78]) is to rely more on meta-analyses, or conduct multi-lab studies for accurate inferences. Another could be the further development of new paradigms that account for the same (or new) confounds, but focus more on detecting reliable individual differences.

When low-level learning is not a "confound"

We recommend that researchers employ diagnostic items as a means to measure adaptive control, because effects on these items are not easily explained by simple response repetition effects or the formation of stimulus-response (or stimulus-stimulus) associations. Therefore, they license consideration of adaptive control accounts. This also led us to label as "confounds" design features that allow for these lower-level associative effects to influence performance, consistent with the decades-old literature on this topic. Importantly, however, this terminology is not meant to imply that stimulus-response or stimulus-stimulus learning is not adaptive in its own right - it most certainly is. In fact, as noted in Box 1, some recent theories have begun to emphasize that adaptive control might rely on the same learning processes as more low-level forms of learning, with the only difference being that they typically operate on more abstract representations (e.g., task difficulty, congruency expectancy, error likelihood). Similarly, we do not want to argue that adaptive control cannot act at the level of specific stimulus or response features. It most likely can, and such specificity could even be considered a marker of its adaptivity rather than a confound. However, demonstrating adaptive control at this level (where it is probably most powerful) comes with the problem that we cannot distinguish it from other low-level explanations (for a partial exception, see the ISPCE).

Of course, in most real-world situations, it is likely that adaptations of cognitive control parameters (like task focus) go hand-in-hand with lower-level learning about the specifics of the environment. This notion is inherent in the study of the CSPCE and ISPCE discussed above, where it is concrete context or stimulus features that condition control processes. Moreover, this has also been investigated in studies that looked at the interaction between lower-level learning and the CSE [22,79–82]. Accordingly, we would like to emphasize that although confound-minimized designs isolate adaptive control processes from lower-level learning processes, we do not view adaptive control as context-free or untethered from lower-level processing. On the contrary, we believe that context and specific stimulus features may be potent drivers of, and interact with, control.

Finally, our recommendations do not imply that there is no value in employing the type of task designs we denote as "confounded" in this article – what constitutes a confound clearly depends on the intention and hypotheses of the researcher. For example, a researcher might have a hypothesis that is agnostic about the mechanism that underlies the effect, have diverging predictions depending on the underlying process (e.g., adaptive control versus contingency learning), or simply be interested in the interaction between these different levels of learning [22,79–82]. Instead, our argument is that if it is a researcher's goal to isolate the effects of adaptive control from concurrent effects of more concrete, lower-level learning processes (as is often the case), then the use of the above-described inducer-diagnostic item design is strongly recommended.

Concluding remarks

The current literature on adaptive control is characterised by a wide heterogeneity of paradigms and designs. Thus, for researchers who are not 'in the weeds' of this field, it can be difficult to infer a consensus view on the steps needed to optimally study adaptive control in conflict tasks. Here, we argue there is one key consideration. Specifically, to measure the hypothetical effects of adaptive control, researchers should employ diagnostic stimuli (and associated responses) that do not overlap with inducer stimuli (and associated responses) that trigger such control (for a partial exception, see the ISPCE).

Of course, creating optimal conditions for measuring adaptive control is not limited to the use of inducer and diagnostic items. For example, researchers should also ensure they have sufficient statistical power, which will depend on the exact conflict task and effect under investigation. It is also worth noting that indices of adaptive control often show a low correlation across, and even within, different conflict tasks (with or without confounds; [83–85]), similar to other measures of cognitive control [86,87]. In other words, while most of these effects can be reliably observed at the group level, they are not optimized for studying interindividual differences. This could also explain why markers of adaptive control often show low or null correlations with other related measures such as working memory capacity (e.g., [88]). In part, this can be expected for tasks that are popular precisely because they produce reliable group effects, and thus low between-subject variance [86,89].

More broadly, although this paper focused on four prominent measures of adaptive control in *conflict tasks* (for other measures in conflict tasks, see Box 4), we believe that the same ideas apply to the study of adaptive control in other tasks. For example, consider studies of task switching. Here, the repetition of task cues [90,91] and stimulus-response features [92] from one trial to the next influences the switch cost, which is thought to reflect, at least partly, the time needed by control processes to reconfigure task sets [93]. Interestingly, similar measures of adaptive control exist in the task switching literature, such as the list-wide [94], context-specific [95], and item-specific [96,97] modulations of the switch cost (for a review, see [98]). However, while some task switching studies have circumvented such low-level repetition confounds by presenting different cues or stimuli on each trial [99], very few have employed confound-minimized designs [92].

In conclusion, the last two decades have witnessed an exponentially increasing interest in the study of adaptive control. During this period, specific indices of adaptive control in conflict tasks and the methods for avoiding experimental confounds have been extensively discussed and revised numerous times. For those outside the field, it can be difficult to follow what is the current optimal means for studying these control processes. In this paper, we presented a consensus view that emphasizes the inclusion of inducer and diagnostic items in the experimental design. We believe that future studies using such 'confound-minimized' designs (in isolation or in combination with other factors, such as contingency learning) will allow researchers to address and revisit exciting research questions (see Outstanding Questions) and substantially advance our understanding of adaptive control.

Acknowledgments

Glossary

Adaptive control	Adaptive control here refers to control processes or executive functions that dynamically adjust processing selectivity in response to changes in the environment or to internal (performance) monitoring signals (e.g., conflicts)
Cognitive control	The term cognitive control, or executive functions, is generally used to describe a set of (not always well-defined) higher-order processes that are thought to direct, correct, and redirect behavior in line with internal goals and current context
(Cognitive) Conflict	Conflict in information processing is thought to occur when two or more mutually incompatible stimulus representations and/or response tendencies are triggered by a stimulus, such as an incongruent stimulus in the Stroop task (invoking, e.g., both "blue" and "red")
Conflict adaptation	Conflict adaptation refers to adaptive processes that are putatively triggered following the detection of conflict, and are recruited for the purpose of resolving this conflict or preventing subsequent occurences of conflict. This term is sometimes also used to refer specifically to the CSE
Contingency learning	Contingency learning refers to the general learning process of forming stimulus-stimulus and/or stimulus-response associations based on their co-occurrence, with the strength of associations increasing as a function of the frequency of co-occurrence. While contingency learning is often discussed within the context of implicit learning, one does not have to make any assumptions about whether this learning occurs explicitly or implicitly, and is strategic or automatic
Diagnostic items	These are the items that are used to "measure" the effects of adaptive control on performance. These items are sometimes also referred to as non-manipulated items, unbiased items, transfer items, or test items
Feature integration	Feature integration refers to the idea that multiple features of a given stimulus are integrated or bound together in perception and memory. An extension of this idea holds that this integration of the features of an experience (or

event) in memory also incorporates one's response to the stimulus into an episodic "event file", and – more recently – that event files might also incorporate internal attentional states. The subsequent presentation of one of those event features is then thought to facilitate the retrieval of the entire event file from memory

Inducer items Researchers manipulate these items to "trigger" adaptive control. These items are sometimes also referred to as manipulated items, biased items, context items, or training items

References

1. Miller EK, & Cohen JD (2001). An integrative theory of prefrontal cortex function. Annual review of neuroscience, 24(1), 167–202.

2. Stroop JR (1935). Studies of interference in serial verbal reactions. Journal of experimental psychology, 18(6), 643.

3. Chiu YC &, Egner T. (2019). Cortical and subcortical contributions to context-control learning. Neuroscience & Biobehavioral Reviews, 99, 33–41. [PubMed: 30685484]

4. Goschke T (2003). Voluntary action and cognitive control from a cognitive neuroscience perspective Voluntary action: Brains, minds, and sociality, 49–85. Oxford University Press.

5. Shenhav A, Botvinick MM, & Cohen JD (2013). The expected value of control: an integrative theory of anterior cingulate cortex function. Neuron, 79(2), 217–240. [PubMed: 23889930]

6. Botvinick MM, Braver TS, Barch DM, Carter CS, & Cohen JD (2001). Conflict monitoring and cognitive control. Psychological review, 108(3), 624. [PubMed: 11488380]

7. Braver TS (2012). The variable nature of cognitive control: a dual mechanisms framework. Trends in cognitive sciences, 16(2), 106–113. [PubMed: 22245618]

8. Kerns JG, Cohen JD, MacDonald AW, Cho RY, Stenger VA, & Carter CS (2004). Anterior cingulate conflict monitoring and adjustments in control. Science, 303(5660), 1023–1026. [PubMed: 14963333]

9. Ridderinkhof KR, Ullsperger M, Crone EA, & Nieuwenhuis S (2004). The role of the medial frontal cortex in cognitive control. Science, 306(5695), 443–447. [PubMed: 15486290]

10. Bugg JM (2014). Evidence for the sparing of reactive cognitive control with age. Psychology and Aging, 29, 115–127. [PubMed: 24378111]

11. Iani C, Stella G, & Rubichi S (2014). Response inhibition and adaptations to response conflict in 6-to 8-year-old children: Evidence from the Simon effect. Attention, Perception, & Psychophysics, 76(4), 1234–1241.

12. Larson MJ, Clawson A, Clayson PE, & South M (2012). Cognitive control and conflict adaptation similarities in children and adults. Developmental Neuropsychology, 37(4), 343–357. [PubMed: 22612546]

13. Abrahamse E, Ruitenberg M, Duthoo W, Sabbe B, Morrens M, & Van Dijck JP (2016). Conflict adaptation in schizophrenia: reviewing past and previewing future efforts. Cognitive neuropsychiatry, 21(3), 197–212. [PubMed: 27100079]

14. Clawson A, Clayson PE, & Larson MJ (2013). Cognitive control adjustments and conflict adaptation in major depressive disorder. Psychophysiology, 50(8), 711–721. [PubMed: 23735120]

15. Lansbergen MM, Kenemans JL, & Van Engeland H (2007). Stroop interference and attention-deficit/hyperactivity disorder: a review and meta-analysis. Neuropsychology, 21(2), 251. [PubMed: 17402825]

16. Larson MJ, South M, Clayson PE, & Clawson A (2012). Cognitive control and conflict adaptation in youth with high-functioning autism. Journal of Child Psychology and Psychiatry, 53(4), 440–448. [PubMed: 22176206]

17. Steudte-Schmiedgen S, Stalder T, Kirschbaum C, Weber F, Hoyer J, & Plessow F (2014). Trauma exposure is associated with increased context-dependent adjustments of cognitive control in patients with posttraumatic stress disorder and healthy controls. Cognitive, Affective, & Behavioral Neuroscience, 14(4), 1310–1319.

18. Tulek B, Atalay NB, Kanat F, & Suerdem M (2013). Attentional control is partially impaired in obstructive sleep apnea syndrome. Journal of sleep research, 22(4), 422–429. [PubMed: 23414228]

19. Wylie SA, Ridderinkhof KR, Bashore TR, & van den Wildenberg WP (2010). The effect of Parkinson's disease on the dynamics of on-line and proactive cognitive control during action selection. Journal of cognitive neuroscience, 22(9), 2058–2073. [PubMed: 19702465]

20. Schmidt JR (2013). Questioning conflict adaptation: proportion congruent and Gratton effects reconsidered. Psychonomic Bulletin & Review, 20(4), 615–630. [PubMed: 23325703]

21. Schmidt JR (2018). Evidence against conflict monitoring and adaptation: An updated review. Psychonomic bulletin & review. In press.

22. Weissman DH, Hawks ZW, & Egner T (2016). Different levels of learning interact to shape the congruency sequence effect. Journal of Experimental Psychology: Learning, Memory, and Cognition, 42(4), 566.

23. Chiu YC, Jiang J, & Egner T (2017). The caudate nucleus mediates learning of stimulus-control state associations. Journal of Neuroscience, 37(4), 1028–1038. [PubMed: 28123033]

24. Jiang J, Brashier NM, & Egner T (2015). Memory meets control in hippocampal and striatal binding of stimuli, responses, and attentional control states. Journal of Neuroscience, 35(44), 14885–14895. [PubMed: 26538657]

25. Egner T (2014). Creatures of habit (and control): a multi-level learning perspective on the modulation of congruency effects. Frontiers in psychology, 5, 1247. [PubMed: 25414679]

26. Bugg JM, & Crump MJ (2012). In support of a distinction between voluntary and stimulus-driven control: A review of the literature on proportion congruent effects. Frontiers in psychology, 3, 367. [PubMed: 23060836]

27. Duthoo W, Abrahamse EL, Braem S, Boehler CN, & Notebaert W (2014). The heterogeneous world of congruency sequence effects: An update. Frontiers in Psychology, 5, 1001. [PubMed: 25250005]

28. Gratton G, Coles MG, & Donchin E (1992). Optimizing the use of information: strategic control of activation of responses. Journal of Experimental Psychology: General, 121(4), 480. [PubMed: 1431740]

29. Egner T (2007). Congruency sequence effects and cognitive control. Cognitive, Affective, & Behavioral Neuroscience, 7(4), 380–390.

30. Mayr U, Awh E, & Laurey P (2003). Conflict adaptation effects in the absence of executive control. Nature neuroscience, 6(5), 450. [PubMed: 12704394]

31. Hommel B, Proctor RW, & Vu KPL (2004). A feature-integration account of sequential effects in the Simon task. Psychological research, 68(1), 1–17. [PubMed: 14752663]

32. Schmidt JR, & De Houwer J (2011). Now you see it, now you don't: Controlling for contingencies and stimulus repetitions eliminates the Gratton effect. Actapsychologica, 138(1), 176–186.

33. Jiménez L, Lupianez J, & Vaquero JMM (2009). Sequential congruency effects in implicit sequence learning. Consciousness and Cognition, 18(3), 690–700. [PubMed: 19464194]

34. Schmidt JR, Crump MJ, Cheesman J, & Besner D (2007). Contingency learning without awareness: Evidence for implicit control. Consciousness and Cognition, 16(2), 421–435. [PubMed: 16899377]

35. Jiménez L, & Méndez A (2014). Even with time, conflict adaptation is not made of expectancies. Frontiers in psychology, 5, 1042. [PubMed: 25278926]

36. Kim S, & Cho YS (2014). Congruency sequence effect without feature integration and contingency learning. Acta psychologica, 149, 60–68. [PubMed: 24704781]

37. Schmidt JR, & Weissman DH (2014). Congruency sequence effects without feature integration or contingency learning confounds. PLoS One, 9(7), e102337. [PubMed: 25019526]

38. Weissman DH, Jiang J, & Egner T (2014). Determinants of congruency sequence effects without learning and memory confounds. Journal of Experimental Psychology: Human Perception and Performance, 40(5), 2022. [PubMed: 25089574]

39. Duthoo W, Abrahamse EL, Braem S, Boehler CN, & Notebaert W (2014). The congruency sequence effect 3.0: a critical test of conflict adaptation. PloS one, 9(10), e110462. [PubMed: 25340396]

40. Mordkoff JT (2012). Observation: Three reasons to avoid having half of the trials be congruent in a four-alternative forced-choice experiment on sequential modulation. Psychonomic Bulletin & Review, 19(4), 750–757. [PubMed: 22549895]

41. Hazeltine E, Lightman E, Schwarb H, & Schumacher EH (2011). The boundaries of sequential modulations: evidence for set-level control. Journal of Experimental Psychology: Human Perception and Performance, 37(6), 1898. [PubMed: 21767054]

42. Kiesel A, Kunde W, & Hoffmann J (2006). Evidence for task-specific resolution of response conflict. Psychonomic Bulletin & Review, 13(5), 800–806. [PubMed: 17328376]

43. Notebaert W, & Verguts T (2008). Cognitive control acts locally. Cognition, 106(2), 1071–1080. [PubMed: 17537419]

44. Braem S, Abrahamse EL, Duthoo W, & Notebaert W (2014). What determines the specificity of conflict adaptation? A review, critical analysis, and proposed synthesis. Frontiers in psychology, 5, 1134. [PubMed: 25339930]

45. Braem S, Verguts T, & Notebaert W (2011). Conflict adaptation by means of associative learning. Journal of Experimental Psychology: Human Perception and Performance, 37(5), 1662. [PubMed: 21728466]

46. Janczyk M, & Leuthold H (2018). Effector system-specific sequential modulations of congruency effects. Psychonomic bulletin & review, 25(3), 1066–1072. [PubMed: 28608004]

47. Lim CE, & Cho YS (2018). Determining the scope of control underlying the congruency sequence effect: roles of stimulus-response mapping and response mode. Acta psychologica, 190, 267–276. [PubMed: 30170247]

48. Bugg JM (2017). Context, conflict, and control. The Wiley handbook of cognitive control, 79–96.

49. Hutchison KA (2011). The interactive effects of listwide control, item-based control, and working memory capacity on Stroop performance. Journal of Experimental Psychology: Learning, Memory, and Cognition, 37(4), 851.

50. Bugg JM (2014). Conflict-triggered top-down control: Default mode, last resort, or no such thing?. Journal of Experimental Psychology: Learning, Memory, and Cognition, 40(2), 567.

51. Bugg JM, & Chanani S (2011). List-wide control is not entirely elusive: Evidence from picture–word Stroop. Psychonomic bulletin & review, 18(5), 930–936. [PubMed: 21638107]

52. Gonthier C, Braver TS, & Bugg JM (2016). Dissociating proactive and reactive control in the Stroop task. Memory & Cognition, 44(5), 778–788. [PubMed: 26861210]

53. Bugg JM, Jacoby LL, & Toth JP (2008). Multiple levels of control in the Stroop task. Memory & cognition, 36(8), 1484–1494. [PubMed: 19015507]

54. Blais C, & Bunge S (2010). Behavioral and neural evidence for item-specific performance monitoring. Journal of Cognitive Neuroscience, 22(12), 2758–2767. [PubMed: 19925177]

55. Schmidt JR (2017). Time-out for conflict monitoring theory: Preventing rhythmic biases eliminates the list-level proportion congruent effect. Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale, 71(1), 52. [PubMed: 27977230]

56. Wühr P, Duthoo W, & Notebaert W (2015). Generalizing attentional control across dimensions and tasks: Evidence from transfer of proportion-congruent effects. The Quarterly Journal of Experimental Psychology, 68(4), 779–801. [PubMed: 25380403]

57. Spinelli G, Perry JR, & Lupker SJ (2019). Adaptation to conflict frequency without contingency and temporal learning: Evidence from the picture-word interference task. Journal of Experimental Psychology: Human Perception & Performance In press.

58. Crump MJC, Gong Z, & Milliken B (2006). The context-specific proportion congruent Stroop effect: Location as a contextual cue. Psychonomic bulletin & review, 13(2), 316–321. [PubMed: 16893001]

59. Lehle C, & Hübner R (2008). On-the-fly adaptation of selectivity in the flanker task. Psychonomic Bulletin & Review, 15(4), 814–818. [PubMed: 18792509]

60. Schouppe N, Ridderinkhof KR, Verguts T, & Notebaert W (2014). Context-specific control and context selection in conflict tasks. Actapsychologica, 146, 63–66.

61. Wendt M, & Kiesel A (2011). Conflict adaptation in time: Foreperiods as contextual cues for attentional adjustment. Psychonomic Bulletin & Review, 18(5), 910–916. [PubMed: 21698510]

62. Bejjani C, Zhang Z, & Egner T (2018). Control by association: Transfer of implicitly primed attentional states across linked stimuli. Psychonomic bulletin & review, 25(2), 617–626. [PubMed: 29450789]

63. Crump MJC, & Milliken B (2009). The flexibility of context-specific control: Evidence for context-driven generalization of item-specific control settings. The Quarterly Journal of Experimental Psychology, 62(8), 1523–1532. [PubMed: 19370487]

64. Crump MJC, Brosowsky NP, & Milliken B (2017). Reproducing the location-based context-specific proportion congruent effect for frequency unbiased items: A reply to Hutcheon and Spieler (2016). The Quarterly Journal of Experimental Psychology, 70(9), 1792–1807. [PubMed: 27340758]

65. Crump MJC, Milliken B, Leboe-McGowan J, Leboe-McGowan L, & Gao X (2018). Context-dependent control of attention capture: Evidence from proportion congruent effects. Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale, 72, 91–104. [PubMed: 29389144]

66. Hutcheon TG, & Spieler DH (2017). Limits on the generalizability of context-driven control. The Quarterly Journal of Experimental Psychology, 70(7), 1292–1304. [PubMed: 27109465]

67. Schmidt JR, & Lemercier C (2019). Context-specific proportion congruent effects: Compound-cue contingency learning in disguise. Quarterly Journal of Experimental Psychology, 72, 1119–1130.

68. Bugg JM, & Hutchison KA (2013). Converging evidence for control of color–word Stroop interference at the item level. Journal of Experimental Psychology: Human Perception and Performance, 39(2), 433. [PubMed: 22845037]

69. Bugg JM, Jacoby LL, & Chanani S (2011). Why it is too early to lose control in accounts of item-specific proportion congruency effects. Journal of Experimental Psychology: Human Perception and Performance, 37(3), 844. [PubMed: 20718569]

70. Bugg JM, & Dey A (2018). When stimulus-driven control settings compete: On the dominance of categories as cues for control. Journal of Experimental Psychology: Human Perception and Performance, 44(12), 1905. [PubMed: 30211592]

71. Jacoby LL, Lindsay DS, & Hessels S (2003). Item-specific control of automatic processes: Stroop process dissociations. Psychonomic Bulletin & Review, 10(3), 638–644. [PubMed: 14620358]

72. Schmidt JR, & Besner D (2008). The Stroop effect: why proportion congruent has nothing to do with congruency and everything to do with contingency. Journal of Experimental Psychology: Learning, Memory, and Cognition, 34(3), 514.

73. Chiu YC, & Egner T (2017). Cueing cognitive flexibility: Item-specific learning of switch readiness. Journal of Experimental Psychology: Human Perception and Performance, 43(12), 1950. [PubMed: 28406686]

74. Schmidt JR, Augustinova M, & De Houwer J (2018). Category learning in the color-word contingency learning paradigm. Psychonomic bulletin & review, 25(2), 658–666. [PubMed: 29404798]

75. Schmidt JR, & De Houwer J (2016). Contingency learning tracks with stimulus-response proportion. Experimental Psychology, 63, 79–88. [PubMed: 27221598]

76. Cronbach LJ, & Furby L (1970). How we should measure" change": Or should we?. Psychological bulletin, 74(1), 68.

77. Crump MJC, & Brosowsky NP (2019). conflictPower: Simulation based power analysis for adaptive control designs. R package version 0.1.0. Retrieved from https://crumplab.github.io/conflictPower/

78. Verbruggen F, Aron AR, Band GP, Beste C, Bissett PG, Brockett AT, [...] & Boehler CN. (2019). A consensus guide to capturing the ability to inhibit actions and impulsive behaviors in the stop-signal task. Elife, 8, e46323. [PubMed: 31033438]

79. Braem S, Hickey C, Duthoo W, & Notebaert W (2014). Reward determines the context-sensitivity of cognitive control. Journal of Experimental Psychology: Human perception and performance, 40(5), 1769. [PubMed: 25089575]

80. Dignath D, Johannsen L, Hommel B, & Kiesel A (in press). Reconciling cognitive-control and episodic-retrieval accounts of sequential conflict modulation: Binding of control-states into event-files. Journal of Experimental Psychology : Human perception and performance.

81. Pires L, Leitão J, Guerrini C, & Simões MR (2018). Cognitive control during a spatial Stroop task: Comparing conflict monitoring and prediction of response-outcome theories. Acta psychologica, 189, 63–75. [PubMed: 28683927]

82. Spapé MM, & Hommel B (2008). He said, she said: Episodic retrieval induces conflict adaptation in an auditory Stroop task. Psychonomic Bulletin & Review, 15(6), 1117–1121. [PubMed: 19001577]

83. Feldman JL, & Freitas AL (2016). An investigation of the reliability and self-regulatory correlates of conflict adaptation. Experimental psychology, 63, 237–247. [PubMed: 27750519]

84. Ruitenberg M, Braem S, Du Cheyne H, & Notebaert W (2019). Learning to be in control involves response-specific mechanisms. Attention, Perception, & Psychophysics. In press.

85. Whitehead PS, Brewer GA, & Blais C (2019). Are cognitive control processes reliable?. Journal of experimental psychology: learning, memory, and cognition, 45, 765–778.

86. Hedge C, Powell G, & Sumner P (2018). The reliability paradox: Why robust cognitive tasks do not produce reliable individual differences. Behavior Research Methods, 50(3), 1166–1186. [PubMed: 28726177]

87. Paap KR, & Sawi O (2016). The role of test-retest reliability in measuring individual and group differences in executive functioning. Journal of Neuroscience Methods, 274, 81–93. [PubMed: 27720867]

88. Meier ME, & Kane MJ (2013). Working memory capacity and Stroop interference: Global versus local indices of executive control. Journal of Experimental Psychology: Learning, Memory, and Cognition, 39(3), 748.

89. Enkavi AZ, Eisenberg IW, Bissett PG, Mazza GL, MacKinnon DP, Marsch LA, & Poldrack RA (2019). Large-scale analysis of test-retest reliabilities of self–regulation measures. Proceedings of the National Academy of Sciences, 116(12), 5472–5477.

90. Logan GD, & Bundesen C (2003). Clever homunculus: Is there an endogenous act of control in the explicit task-cuing procedure?. Journal of Experimental Psychology: Human Perception and Performance, 29(3), 575. [PubMed: 12848327]

91. Mayr U, & Kliegl R (2000). Task-set switching and long-term memory retrieval. Journal of Experimental Psychology: Learning, Memory, and Cognition, 26(5), 1124–1140.

92. Schmidt JR, & Liefooghe B (2016). Feature integration and task switching: Diminished switch costs after controlling for stimulus, response, and cue repetitions. PloS one, 11(3), e0151188. [PubMed: 26964102]

93. Monsell S (2003). Task switching. Trends in cognitive sciences, 7(3), 134–140. [PubMed: 12639695]

94. Dreisbach G, & Haider H (2006). Preparatory adjustment of cognitive control in the task switching paradigm. Psychonomic Bulletin & Review, 13(2), 334–338. [PubMed: 16893004]

95. Crump MJC, & Logan GD (2010). Contextual control over task-set retrieval. Attention, Perception, & Psychophysics, 72(8), 2047–2053.

96. Chiu YC, & Egner T (2017). Cueing cognitive flexibility: Item-specific learning of switch readiness. Journal of Experimental Psychology: Human Perception and Performance, 43(12), 1950. [PubMed: 28406686]

97. Leboe JP, Wong J, Crump M, & Stobbe K (2008). Probe-specific proportion task repetition effects on switching costs. Perception & Psychophysics, 70(6), 935–945. [PubMed: 18717381]

98. Braem S, & Egner T (2018). Getting a grip on cognitive flexibility. Current directions in psychological science, 27(6), 470–476. [PubMed: 30555214]

99. Schneider DW (2015). Isolating a mediated route for response congruency effects in task switching. Journal of Experimental Psychology: Learning, Memory, and Cognition, 41(1), 235.

100. Abrahamse E, Braem S, Notebaert W, & Verguts T (2016). Grounding cognitive control in associative learning. Psychological Bulletin, 142(7), 693. [PubMed: 27148628]

101. Blais C, Robidoux S, Risko EF, & Besner D (2007). Item-specific adaptation and the conflict monitoring hypothesis: a computational model. Psychological Review, 114, 1076–1086. [PubMed: 17907873]

102. Verguts T, & Notebaert W (2008). Hebbian learning of cognitive control: dealing with specific and nonspecific adaptation. Psychological review, 115(2), 518. [PubMed: 18426302]

103. Egner T (2008). Multiple conflict-driven control mechanisms in the human brain. Trends in cognitive sciences, 12(10), 374–380. [PubMed: 18760657]

104. Botvinick MM (2007). Conflict monitoring and decision making: reconciling two perspectives on anterior cingulate function. Cognitive, Affective, & Behavioral Neuroscience, 7(4), 356–366.

105. Alexander WH, & Brown JW (2011). Medial prefrontal cortex as an action-outcome predictor. Nature neuroscience, 14(10), 1338. [PubMed: 21926982]

106. Silvetti M, Seurinck R, & Verguts T (2011). Value and prediction error in medial frontal cortex: integrating the single-unit and systems levels of analysis. Frontiers in human neuroscience, 5, 75. [PubMed: 21886616]

107. Ridderinkhof RK (2002). Micro-and macro-adjustments of task set: activation and suppression in conflict tasks. Psychological research, 66(4), 312–323. [PubMed: 12466928]

108. Weissman DH, Colter KM, Grant LD, & Bissett PG (2017). Identifying stimuli that cue multiple responses triggers the congruency sequence effect independent of response conflict. Journal of Experimental Psychology: Human Perception and Performance, 43(4), 677. [PubMed: 28095005]

109. Schmidt JR (2013). Temporal learning and list-level proportion congruency: Conflict adaptation or learning when to respond?. PLoS One, 8(11), e82320. [PubMed: 24312413]

110. Schmidt JR, & Weissman DH (2016). Congruency sequence effects and previous response times: Conflict adaptation or temporal learning?. Psychological Research, 80(4), 590–607. [PubMed: 26093801]

111. Cohen-Shikora ER, Suh J, & Bugg JM (2018). Assessing the temporal learning account of the list-wide proportion congruence effect. Journal of Experimental Psychology: Learning, Memory, and Cognition. In press.

112. Logan GD, & Zbrodoff NJ (1979). When it helps to be misled: Facilitative effects of increasing the frequency of conflicting stimuli in a Stroop-like task. Memory & cognition, 7(3), 166–174.

113. Aben B, Calderon CB, Van der Cruyssen L, Picksak D, Van den Bussche E, & Verguts T (2019). Context-dependent modulation of cognitive control involves different temporal profiles of fronto-parietal activity. NeuroImage, 189, 755–762. [PubMed: 30735827]

114. De Pisapia N, & Braver TS (2006). A model of dual control mechanisms through anterior cingulate and prefrontal cortex interactions. Neurocomputing, 69(10–12), 1322–1326.

115. Aben B, Verguts T, & Van den Bussche E (2017). Beyond trial-by-trial adaptation: A quantification of the time scale of cognitive control. Journal of Experimental Psychology: Human Perception and Performance, 43(3), 509. [PubMed: 28080112]

116. Jiang J, Heller K, & Egner T (2014). Bayesian modeling of flexible cognitive control. Neuroscience & Biobehavioral Reviews, 46, 30–43. [PubMed: 24929218]

117. Torres-Quesada M, Lupiáñez J, Milliken B, & Funes MJ (2014). Gradual proportion congruent effects in the absence of sequential congruent effects. Acta Psychologica, 149, 78–86. [PubMed: 24742686]

118. Funes MJ, Lupiáñez J, & Humphreys G (2010). Sustained vs. transient cognitive control: Evidence of a behavioral dissociation. Cognition, 114(3), 338–347. [PubMed: 19962136]

119. Torres-Quesada M, Funes MJ, & Lupiáñez J (2013). Dissociating proportion congruent and conflict adaptation effects in a Simon-Stroop procedure. Acta psychologica, 142(2), 203–210. [PubMed: 23337083]

120. Kornblum S, Hasbroucq T, & Osman A (1990). Dimensional overlap: cognitive basis for stimulus-response compatibility--a model and taxonomy. Psychological review, 97(2), 253. [PubMed: 2186425]

121. Egner T, Etkin A, Gale S, & Hirsch J (2007). Dissociable neural systems resolve conflict from emotional versus nonemotional distracters. Cerebral cortex, 18(6), 1475–1484. [PubMed: 17940084]

122. Bombeke K, Langford ZD, Notebaert W, & Boehler CN (2017). The role of temporal predictability for early attentional adjustments after conflict. PloS one, 12(4), e0175694. [PubMed: 28410395]

123. Schuch S, Dignath D, Steinhauser M, & Janczyk M (2018). Monitoring and control in multitasking. Psychonomic bulletin & review. In press.

124. Goschke T (2000). Intentional reconfiguration and J-TI Involuntary Persistence In Task Set Switching. Control of cognitive processes: Attention and performance XVIII, 18, 331.

125. Brown JW, Reynolds JR, & Braver TS (2007). A computational model of fractionated conflict-control mechanisms in task-switching. Cognitive psychology, 55(1), 37–85. [PubMed: 17078941]

126. Braem S, Verguts T, Roggeman C, & Notebaert W (2012). Reward modulates adaptations to conflict. Cognition, 125(2), 324–332. [PubMed: 22892279]

127. Arrington CM, & Logan GD (2004). The cost of a voluntary task switch. Psychological science, 15(9), 610–615. [PubMed: 15327632]

128. Orr JM, Carp J, & Weissman DH (2012). The influence of response conflict on voluntary task switching: A novel test of the conflict monitoring model. Psychological Research, 76(1), 60–73. [PubMed: 21380567]

129. Dignath D, Kiesel A, & Eder AB (2015). Flexible conflict management: conflict avoidance and conflict adjustment in reactive cognitive control. Journal of Experimental Psychology: Learning, Memory, and Cognition, 41(4), 975.

130. Desender K, Buc Calderon C, Van Opstal F, & Van den Bussche E (2017). Avoiding the conflict: Metacognitive awareness drives the selection of low-demand contexts. Journal of Experimental Psychology: Human Perception and Performance, 43(7), 1397. [PubMed: 28368164]

Outstanding Questions Box

- Will previous findings hold when using the recommended measures of adaptive control? For example, topics such as the domain-generality of control processes, their reward-sensitivity, the role of consciousness, or the relation to other constructs such as working memory capacity, have mostly been studied using designs that potentially measured stimulus-response or stimulus-stimulus learning.

- Are the learning processes mediating adaptive control qualitatively different from, or similar to, the learning processes that mediate stimulus-response or stimulus-stimulus learning?

- What learning processes support transfer and generalization of control within and across contexts and items?

- Which brain regions or processes are involved in adaptive control when using inducer-diagnostic item designs, and how do they relate to the neural mechanisms that underlie stimulus-response or stimulus-stimulus learning?

- What role, if any, does the detection of conflict play in driving adaptive control? For example, is there a functional role for conflict in triggering adaptive control processes?

- Can future designs be further modified to detect more reliable individual differences within and across tasks? And what are the correlates of individual differences in adaptive control?

Trends Box

- Early putative indices of adaptive control in conflict tasks spurred a great deal of research, but also numerous discussions on what these indices actually measure.

- Recent studies have shown that adaptive control effects can be observed after controlling for low-level confounds. However, many canonical findings in the literature, for instance concerning the functional neuroanatomy of adaptive control, are based on older, confounded designs, and may thus be subject to revision.

- This research field is now starting to experience a second wave of studies on adaptive control in conflict tasks with improved designs that allow us to (re)address old and new questions.

Box 1.

Cognitive control versus low-level learning

The observation that low-level learning contributes to various metrics of adaptive control has led to several theoretical discussions as to how adaptive control works, and how 'low-level' stimulus-stimulus or stimulus-response learning relates to cognitive control more broadly. Initially, the seminal conflict monitoring account (Figure IA;[6]) proposed that conflict is detected by a conflict monitor, which signals the need for adaptation to specific top-down task or control modules. However, some have subsequently argued that there is no need for such modules to explain adaptive control (Figure IC; [21]), while others have suggested that these top-down modules are recruited as a last resort mechanism [50]. Yet other theories emphasize a close integration between low-level learning and cognitive control (Figure IB, D; [25,100–102]) to better account for observations such as context- and item-specific proportion congruency effects, or the domain-specificity of the congruency sequence effect (for reviews, see [44,103]). These theories suggest that control representations are embedded in the same associative network of stimulus and response representations. Some see an important role for conflict as a teaching signal that promotes these different forms of learning (**B;** [101,102]), while others have been agnostic about this [25], or argued against conflict as a teaching signal in driving adaptive control (**D;**[22]).

Taken together, we believe these theories differ in their response to two pertinent questions. First, some accounts consider cognitive control and low-level learning to rely on separate mechanisms or modules (**A,C**), while others question this idea and capitalize on the similarities between the two by arguing for the associative learning or episodic binding of control states (**B,D**). A second open question has been whether conflict serves as an active control signal (**A,B**), or plays no necessary functional role (**C,D**) in adaptive control.

Finally, two more issues are important to mention. First, there also exist various more specific theories on the nature of the proposed conflict monitor [5,104–106], and the precise mechanisms that make up these control processes ([28,107], for an overview, see [108]), that are beyond the scope of this article. Second, while the different measures discussed in this article are all thought to measure adaptive control beyond stimulus-response learning, they can still differ in their 'degree' of higher-order processing. For example, one interesting proposal is that adaptive control processes can also come about by simply learning one abstract property of the task, namely, the time it takes to respond to the previous trial (i.e., temporal learning accounts; [109,110]; but see [57,111]).

Figure I.
An abstract schematic depiction of four major groups of theories on adaptive control.
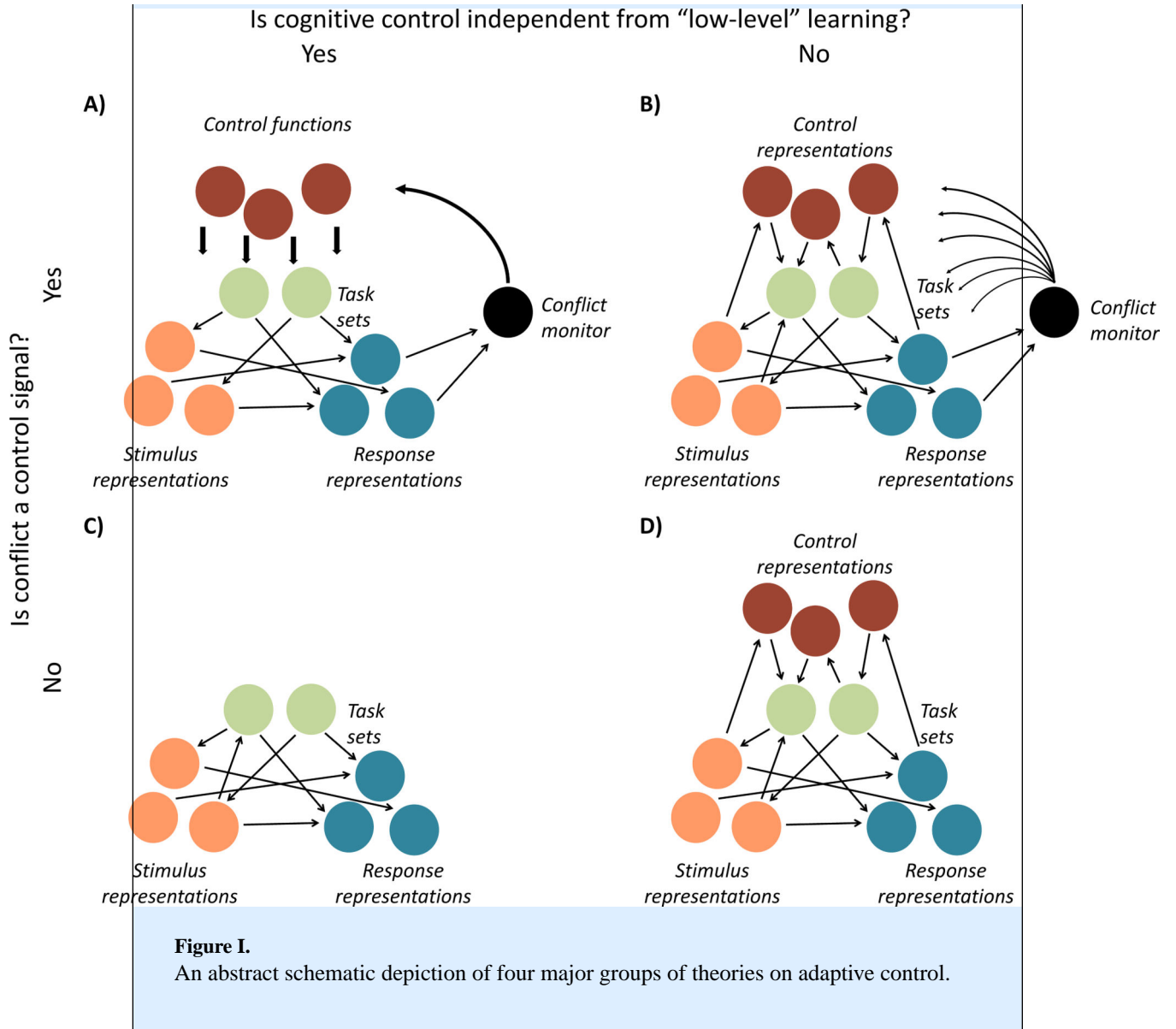
Box 2.

Different measures of adaptive control – What do they test?

The most widely used marker of adaptive control in conflict tasks is the congruency sequence effect (CSE). The CSE refers to the observation that the congruency effect is smaller following incongruent as compared to congruent trials (Figures 1A and 2; [28]; for reviews, see [27,29]). The CSE is typically thought to measure a short-lived, reactive adaptation to a just-experienced conflict between competing response representations [6].

A second marker is the list-wide proportion congruency effect (LWPCE). The LWPCE refers to the observation that the congruency effect is smaller in blocks that contain mostly incongruent trials than in blocks that contain mostly congruent trials (Figure 1B; Figure 2; [28,112]; for a review, see [26]). In contrast to the CSE, the LWPCE is usually considered to measure more global, and possibly proactive (i.e., anticipatory; [113,114]), adaptations of control that accrue over a larger time-window than the previous trial. However, some researchers have suggested that these two effects might be mediated by the same underlying learning processes [6,35,102,115,116]. In fact, one can even consider the typical LWPCE design as being confounded with different proportions of previous-trial congruency, according to which the LWPCE could also reflect an accumulation of CSE effects. While some studies have accounted for this, and still found a small LWPCE effect [117], the exact relation between these two types of adaptive control is still a topic of discussion.

For anyone new to the field who is interested in assessing adaptive control, we recommend focusing on the CSE or LWPCE, as these effects have been studied most extensively with the here-described confound-minimized designs.

A third and fourth way to manipulate proportion congruency is to link it to specific contexts or specific items, respectively. In tasks measuring the context-specific proportion congruency effect (CSPCE; [58]), the proportions of congruent and incongruent trials are tied to a contextual feature that can change on a trial-by-trial basis and is not part of the task stimulus itself (e.g., the spatial location of Stroop stimuli; Figure 1C). In tasks measuring the item-specific proportion congruency effect (ISPCE; [71]), the proportions of congruent and incongruent trials are contingent on a specific *task-relevant* feature (e.g., color in the Stroop task; Figure 1D). Similar to the LWPCE, both the CSPCE and the ISPCE are thought to reflect adaptive control processes that learn about proportion congruency across several trials, rather than just the previous trial as for the CSE. In contrast to the LWPCE and more like the CSE, however, these adaptive control processes are recruited reactively after stimulus onset in the CSPCE and ISPCE.

Importantly, please note that all original observations of these four effects came from designs that are now considered confounded, because low-level learning (e.g., contingency learning, associative learning, episodic memory of stimulus-response episodes) could explain the modulation of the congruency effect.

Box 3.

Different conflict tasks – Which one should I choose?

We use the Stroop task [2] as an example throughout our paper, as it is the most popular of all conflict tasks. However, researchers have also used other paradigms to study adaptive control. For instance, in the flanker task participants identify a central target while ignoring congruent or incongruent flanking distractors. As another example, in the Simon task participants make a lateralized response (e.g., a left button press) to identify a non-spatial feature of a stimulus (e.g., its color - red) while ignoring the location of the stimulus on the screen, which can be congruent (e.g., left) or incongruent (e.g., right).

The same signatures of adaptive control (like the CSE, LWPCE, etc.) have been found across different types of conflict tasks. However, there is currently little evidence for adaptive control crossing between conflict tasks for the CSE (i.e., where the inducer trial is from one task, and the diagnostic from another; for reviews, see [44,103]), and only some for the LWPCE [56,118,119]. Further, indices of adaptive control within different conflict tasks show low to no correlations with one another [85]. Thus, one should think carefully about which conflict task to employ. In so doing, at least two things should be considered: First, researchers posit that different tasks invoke different types of conflict (and adaptive control processes) as a function of the cognitive processing stage(s) at which an overlap occurs between task-relevant and task-irrelevant stimulus features [120]. For instance, whereas conflict in the Stroop task likely involves clashing semantic features, conflict in the Simon task likely involves location-triggered response priming. One may also consider employing tasks that invoke other types of conflict, such as emotional conflict [121].
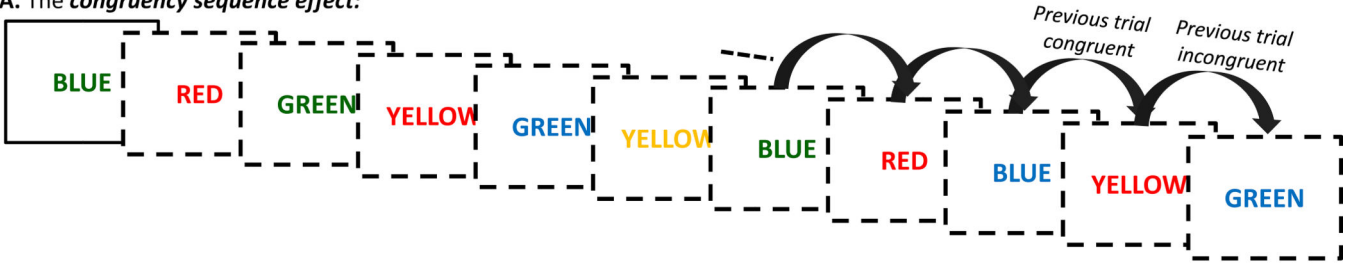
Second, different conflict tasks may have different degrees of power to detect changes in congruency effects. For example, while the CSE can be reliably observed in the Stroop task ([35,85], but see [38]), but its effect size is relatively small (e.g., N=178, ηp^2=05, in [85]). In contrast, tasks wherein the distractor precedes the target, such as the prime-probe task, produce a CSE that has a larger effect size [37]. Similarly, Stroop or flanker tasks wherein the task-irrelevant information appears before target presentation also result in larger CSE effects [122]. Therefore, we strongly recommend running a power analysis before setting up an experiment and choosing the sample size and number of trials accordingly (main text).

Box 4.

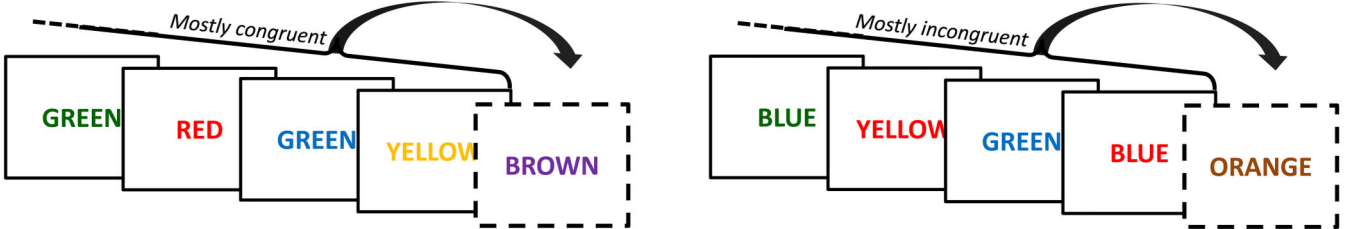### More measures of adaptive control in conflict tasks

This paper focuses on four popular indices of adaptive control in conflict tasks, but other interesting measures exist as well, especially when using variations of conflict tasks such as conflict task switching studies, or free-choice paradigms. First, while most indices of adaptive control come from single task paradigms, some can also be observed in (conflict) task-switching paradigms (for a review, see [123]). A seminal finding in this domain is that the switch cost (i.e., the performance cost associated with alternating versus repeating the just-performed task) is greater when the previous trial was incongruent as compared to congruent [124]. This effect is thought to measure enhanced processing of task-relevant information after an incongruent trial due to adaptive control processes (e.g., [125,126]; but see [92]), similar to some explanations of the CSE [6].

A second group of studies has begun to look at adaptations to conflict in free-choice tasks. All measures discussed thus far come from tasks wherein participants are clearly instructed to perform one task or the other. However, in an attempt to create more ecologically valid paradigms, some researchers have started to employ voluntary task switching paradigms, wherein participants choose which task to perform in each trial [127]. Using such paradigms, some researchers have studied whether people are more likely to repeat a task, or switch to a new task, if the previous trial was incongruent as compared to congruent [128,129]. Similarly, others have studied whether people, when given a choice, tend to avoid or approach contexts that are associated with a higher proportion of incongruent trials [60,130]. Many of these studies have demonstrated that people tend to prefer congruent over incongruent trials (and associated tasks), which is consistent with the idea that adaptive control is also costly and demanding [5].
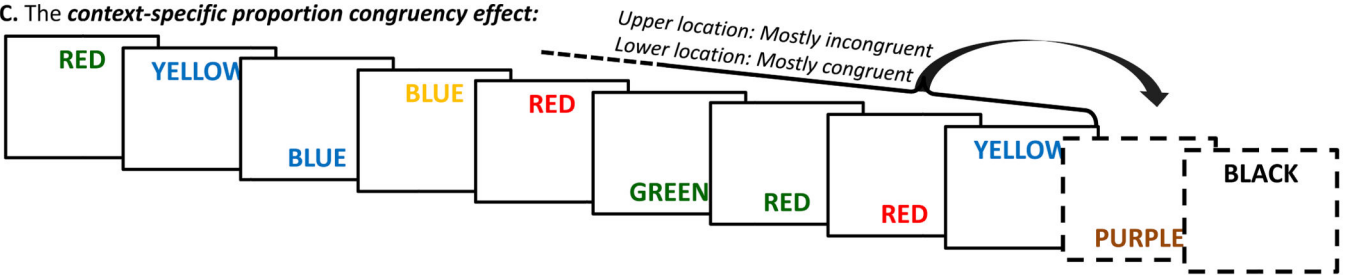
A. The *congruency sequence effect:*

Previous trial congruent

Previous trial incongruent

BLUE RED GREEN YELLOW GREEN YELLOW BLUE RED BLUE YELLOW GREEN

B. The *list-wide proportion congruency effect:*

Mostly congruent

GREEN RED GREEN YELLOW BROWN

Mostly incongruent

BLUE YELLOW GREEN BLUE ORANGE

C. The *context-specific proportion congruency effect:*

Upper location: Mostly incongruent
Lower location: Mostly congruent

RED YELLOW BLUE BLUE RED GREEN RED RED YELLOW PURPLE BLACK

D. The *item-specific proportion congruency effect:*

Red or yellow font: Mostly incongruent
Green or blue font: Mostly congruent

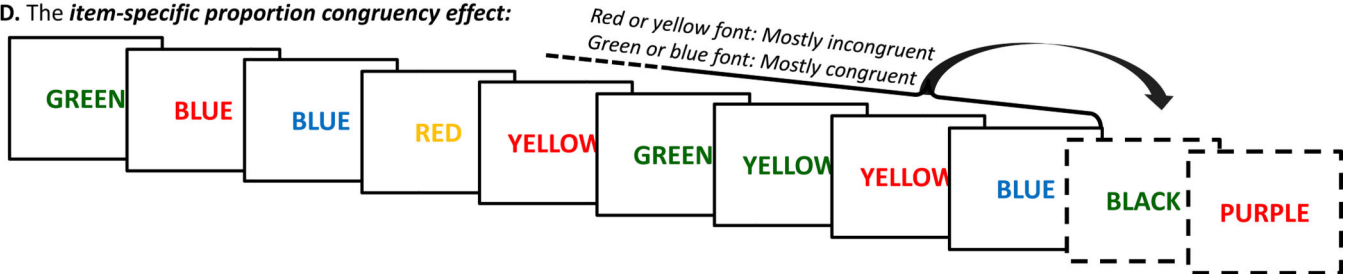GREEN BLUE BLUE RED YELLOW GREEN YELLOW YELLOW BLUE BLACK PURPLE

Figure 1. Different indices of adaptive control in the Stroop task.
For each measure, dashed screens indicate the use of diagnostic items to study adaptive changes in behavior that were triggered by different items. Note that for the examples B-D diagnostic items are usually randomly presented among the other items (not only at the end of a run), which is here not the case for reasons of figure space. A. The congruency sequence effect measures adaptive changes following incongruent versus congruent trials. B. The proportion congruency effect measures adaptations in control in blocks where mostly congruent trials are presented (left) versus blocks where mostly incongruent trials are presented (right). C. The context-specific proportion congruency effect investigates adaptive control in contexts associated with mostly incongruent items (e.g., stimuli presented at the top of the screen) versus contexts with mostly congruent items (e.g., stimuli presented at the bottom of the screen). D. The item-specific proportion congruency effect probes adaptations in control to item features that are presented under mostly incongruent conditions (i.e., the

ink colors red and yellow) versus item features presented in mostly congruent conditions (i.e., the ink colors blue and green).

previous trial
congruent

previous trial
incongruent

= congruency sequence effect

mostly congruent
block

mostly incongruent
block

= list-wide PCE

mostly congruent
context

mostly incongruent
context

= context-specific PCE

mostly congruent
item

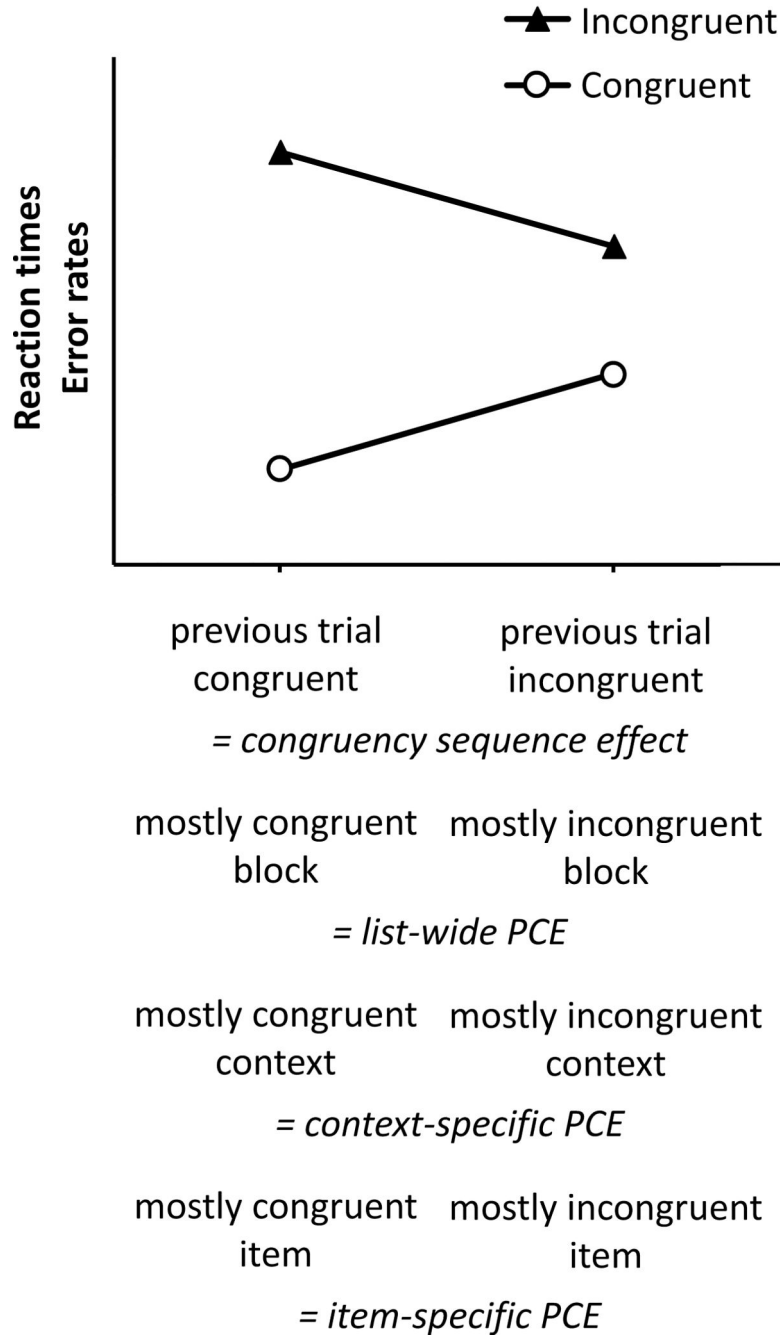mostly incongruent
item

= item-specific PCE

Figure 2. Behavioral indices of adaptive control.
The four performance indices are characterized by a reduction in the congruency effect
either following incongruent versus congruent trials (i.e., congruency sequence effect), or in
mostly incongruent versus mostly congruent conditions (i.e., proportion congrueny effects;
PCEs). The interaction depicted in this figure represents a generic form—the actual form of
the interaction may vary from one index to another